

SEQUENCE LISTING

<110> KIM, Young Tae
LEE, Jae Hyung

<120> Gene involved in the biosynthesis of carotenoid and marine
microorganism, paracoccus haeundaesis, producing the
carotenoid

<130> 428.1056

<150> PCT/KR2004/000752
<151> 2003-03-31

<150> KR2003-20222
<151> 2003-03-31

<150> KR2003-20023
<151> 2003-03-31

<160> 18

<170> KopatentIn 1.71

<210> 1
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer for Paracoccus haeundaesis 16S rDNA

<400> 1
cataagtaat tatggttttg t
21

<210> 2
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse primer for Paracoccus haeundaesis 16S rDNA

<400> 2
cgcttcctta gaaaggag
18

<210> 3
<211> 1454
<212> DNA

<213> Paracoccus haeundaesis

<400> 3

caacttgaga gtttgatcct ggctcagaac gaacgctggc ggcaggctta acacatgcaa
60

gtcgagcgag accttcgggt ctagcggcgg acgggtgagt aacgctggg aacgtgccct
120

tctctacgga atagccccgg gaaactggga gtaataccgt atacgccctt tgggggaaag
180

atttatcgga gaaggatcgg cccgcgttgg attaggtagt tgggtgggta atggcccacc
240

aagccgacga tccatagctg gtttgagagg atgatacagg aactggggac tgagacacgg
300

cccagactcc tacgggaggc agcagtgggg aatcttagac aatgggggca accctgatct
360

agccatgccg cgtgagtgat gaaggcctta gggttgtaa gctctttcag ctgggaagat
420

aatgacggta ccagcagaag aagccccggc taactccgtg ccagcagccg cggtaatagc
480

gagggggcta gcgttggtcg gaattactgg gcgtaaagcg cacgtgggcg gactggaaag
540

tcagaggtga aatcccaggg ctcaaccttg gaactgcctt tgaaactatc agtctggagt
600

tcgagagagg tgagtggaaat tccgagtgtg gaggtgaaat tcgtagatat tcggaggaac
660

accagtggcg aaggcggctc actggctcga tactgacgct gaggtgcaa agcgtgggga
720

gcaaacagga ttagataccc tggtagtcca cgccgtaaac gatgaatgcc agacgtcggc
780

aagcatgctt gtcggtgtca cacctaacgg attaagcatt ccgcctgggg agtacggtcg
840

caagattaaa actcaaagga attgacgggg gcccgacaaa gcggtggagc atgtgggtta
900

attcgaagca acgcgcagaa ccttaccaac ccttgacatg gcaggaccgc tggagagatt
960

cagctttctc gtaagagacc tgcaacacagg tgctgcatgg ctgtcgtcag ctcgtgtcgt
1020

gagatgttcg gttaagtccg gcaacgagcg caaccacgt ccctagttgc cagcattcag
1080

ttgggcactc tatggaaact gccgatgata agtcggagga aggtgtggat gacgtcaagt

1140

tctcatggcc cttacgggtt gggctacaca cgtgctacaa tggtggtgac agtgggttaa
1200

tccccaaaag ccattctcagt tcggattgtc ctctgcaact cgagggcatg aagttggaat
1260

cgctagtaat cgcggaacag catgccgcgg tgaatacgtt cccgggcctt gtacacaccg
1320

cccgtcacac catgggagtt ggttctaccc gacgacgctg cgctaacctt cggggggcag
1380

gcggccacgg taggatcagc gactgggggtg aagtcgtaac aaggtagccg taggggaacc
1440

tgcggtgga tcac
1454

<210> 4
<211> 6223
<212> DNA
<213> crt gene

<400> 4
gttccacgac tggggcatcc ccacgaccgc gtcgctgcgc gccatcgcg ccatgatggg
60

gccggaccgg gttctggtcg ggtcggggcg ggtgcgtcac gggctggacg ccgcgcgggc
120

catccgcctc ggcgcgacc tcgtggggca ggcgccccgc gcgctgccg ccgcgcgcca
180

cagcgccgag gccctgtccg atcacctgtc cgacgtcgtg acccagctgc gcatcgcgat
240

gttctgcacc ggatcgggcg accttgacg gctgcgtgc gcgcctctgc tggtgccggg
300

gccgggtggc caatggtcgc aagcaacggg gatggaaacc ggcatgcgg gactgtagtc
360

tgcgcggtac gccggtccgg gggacaagat gagcgacat gccctgccca aggcagatct
420

gaccgccacc agcctgatcg tctcggggcg catcatcgcc gcgtgggtgg ccctgcatgt
480

gcatcgctg tggtttctgg acgcggcggc gcatcccatc ctggcgatcg cgaatttctt
540

ggggctgacc tggctgtcgg tcggtctgtt cttcatcgcg catgacgca tgcacgggtc
600

ggtcgtgccg gggcgctccg gcggcaatgc ggcatgggc cagctggtcc tgtggctgta

660

tgccggattt tcgtggcgca agatgatcgt caagcacatg gcccatcacc gccataccgg
720

aaccgacgac gaccccattt tcgaccatgg cggcccggtc cgctggtacg cgcgcttcat
780

cggcacctat ttcggctggc gcgaggggct gctgctgcc gtcacgtga cggctctatgc
840

gctgatcctg ggggatcgtt ggatgtacgt ggtcttctgg ccgctgccgt cgatcctggc
900

gtcgatccag ctgttcgtgt tcggcacctg gctgccgac cggcccggcc acgacgcgtt
960

ccgggaccgc cataatgcgc ggtcgtcgcg gatcagcgc cccgtgtcgc tgctgacctg
1020

ctttcacttt ggtggttatc atcacgaaca ccacctgcac ccgacggtgc cttggtggcg
1080

cctgccagc acccgaccca agggggacac cgcacgacca atttctgat cgtcgtcgcc
1140

accgtgctgg tgatggagtt gacggcctat tccgtccacc gttggatcat gcacggcccc
1200

ctgggctggg gctggcaca gtcccaccac gaggaacacg accacgcgtt ggaaaagaac
1260

gacctgtacg gcctggtctt tcggtgatc gccacggtgc tgttcacggt gggctggatc
1320

tgggcgccgg tcctgtggtg gatcgctttg ggcatgaccg tctatgggct gatctatttc
1380

gtcctgcatg acgggctggt tcacacgcgc tggccgttcc gctatatccc gcgcaagggc
1440

tatgcccgcc gcctgtatca ggcccaccgc ctgcaccacg cggtcgaggg acgcgacat
1500

tgcgtcagct tcggcttcat ctatgcgcgc ccggtcgaca agctgaagca ggacctgaag
1560

acgtcggggc tgctgcgggc cgaggcgacg gagcgacgt gacctatgac gtgctgctgg
1620

caggggcggg ccttgcaaac gggctgatcg ccctggcgct gcgcgcggcg cggcccgaac
1680

tgccgggtgct gctgctggat catgcggcg gaccgtcaga cggccatacc tggctctgcc
1740

acgaccccca tctgtcgccg cactggctgg cgcggctgaa gccctgcgc cgcgccaact
1800

ggcccgaacca ggaggtgcgc tttccccgcc atgcccggcg gctggccacc ggttacgggt
1860
cgctggacgg ggcggcgctg gcggatgcgg tggcccggtc gggcgccgag atccgctgga 1920
acagcgacat cgccctgctg gatgaacagg gggcgacgct gtcttgccgc acccggatcg
1980
aggcggggcg ggtcctggac gggcgcgggc cgacgcccgc gcggcatctg accgtgggtt
2040
tccagaaatt cgtggcgctc gagatcgaga ccgactgcc ccacggcgctg ccccgccga
2100
tgatcatgga cgcgaccgtc acccagcagg acgggtaccg attcatctat ctgctgcct
2160
tctctccgac gcgcatcctg atcgaggaca ctcgctattc cgatggcggc aatctggacg
2220
acgacgcgct ggcggcgggc tcccacgact atgcccgcc gcagggctgg accggggcg
2280
aggtcggcg cgaacgcggc atcctgccc ttgcgctggc ccatgacgc gcgggcttct
2340
gggcccgatca cgcggagggg cctgttccc tgggactgcg cgcgggggtc tttcaccgg
2400
tcaccggcta ttcgctgcc tatgcggcg aggtggcgga cgtggtggcg ggcctgtccg
2460
ggccgcccgg caccgacgc ctgcgcggcg ccatccgga ttacgcgatc gaccgggac
2520
gccgtgaccg ctttctgcgc ctgctgaacc ggatgctgtt ccgcggctgc gcgcccgacc
2580
ggcgtatac cctgctgcag cggttctacc gcatgcgca tggactgatc gaacggttct
2640
atgcccggcg gctgagcgtg gcggatcagc tgcgcatcgt gaccggcaag cctcccattc
2700
cccttggcac ggccatccgc tgctgcccg aacgtcccct gctgaaggaa aacgcatgaa
2760
cgccattcg cccgcggcca agaccgcat cgtgatcggc gcaggcttg gcgggctggc
2820
cctggccatc cgcctgcagt ccgcgggcat cgccaccacc ctggtcgagg cccgggacaa
2880
gccggcggg cgccctatg tctggcacga tcagggccat gtcttcgacg cgggcccgc
2940
cgtcatcacc gacccgatg cgtcaagga gctgtggcg ctgaccggg aggacatggc

3000

gcgcgacgtg acgctgatgc cgggtgtcgcc cttctatcga ctgatgtggc cgggcgggaa
3060

ggtcttcgat tacgtgaacg aggccgatca gctggagcgc cagatcgccc agttcaaccc
3120

ggacgacctg gaaggatacc gccgcttcg tgattacgcg gaggaggtgt atcaggaggg
3180

ctacgtcaag ctgggcaccg tgccttcct caagctgggc cagatgctca aggccgcgcc
3240

cgcgctgatg aagctggagg cctataagtc cgtccatgcc aaggtcgcg ccttcatcaa
3300

ggacccctat ctgcggcagg cgttttcgta tcacacgctg ctggtgggcg ggaatccctt
3360

ctcgaccagc tcgatctatg cgctgatcca cgcgctggag cggcgcggcg gggctctggtt
3420

cgccaagggc ggaccaacc agctggtcgc gggcatggtc gcgctgttcg aacggcttgg
3480

cggccagatg atgctgaacg ccaaggtcgc ccgcatcgag accgagggcg cgcggaccac
3540

gggcgtcacc ctggcggacg ggcggctctt aagggccgac atggtcgcca gcaacggcga
3600

cgtcatgcac aactatcgcg acctgctggg ccacacggcc cgcgggcaga gccgcgcgaa
3660

atcgctggac cgcaagcgct ggtccatgtc gttgttcgtg ctgcatttcg gtctgcgcga
3720

ggcgcccaag gacatcgcg atcacaccat cctgttcggc ccccgctaca gggagctggt
3780

caacgagatc ttcaagggcc cgaagctggc cgaggatttc tcgctgtacc tgcattcgcc
3840

ctgcacgacc gatccggaca tggcgctcc gggcatgtcc acgcattacg tgctggcccc
3900

cgtgccgcat ctgggccgcg ccgagatcga ttgggcggtc gaggggccgc gctatgccga
3960

ccgcatcctg gcgtccctgg aggagcggct gatcccgaa ctgcgcgcca acctgaccac
4020

gacgcgcatc ttcacgcccg ccgatttcgc cagcgaactg aacgcccac acggcagcgc
4080

cttctcggtc gagccgatcc tgacgcaatc cgcgtgggtc cggccgcaca accgcgacaa
4140

gacgatccgc aacttctatc tggtcggcgc gggcacccat ccgggcgcgc gcattccggg
4200
cgtcgtgggc tcggccaagg ccacggccca ggtgatgctg tccgacctg cgggcgcatg
4260
agcgatctgg tcctgacctc gaccgaggcg atcacccaag ggtcgcaaag ctttgccacg
4320
gcggccaagc tgatgccgcc gggcatccgc gacgacacgg tgatgctcta tgcctggtgc
4380
cgccacgcgg atgacgtgat cgacggtcag gccctgggca gccgccccga ggcggtgaac
4440
gacccgcagg cgcggttga cggcctgcgc gtcgacacgc tggcgccct gcagggcgac
4500
ggtccggtga ccccgccctt tgccgcgctg cgcgcggtg cgcggcgga tgatttccc
4560
caggcctggc ccatggacct gatcgaaggc ttcgcgatg atgtcgagg gcgcgactat
4620
cgcacgtgg atgacgtgct ggaatattcc tatcacgtc caggcatcgt cgcggtgatg
4680
atggccgcgc tgatgggcgt gcgcgacgat cctgtcctgg accgcgcctg cgacctggg
4740
ctggcgttcc agctgaccaa catcgcgcgc gacgtgatc acgatgcgc catcgggcgg
4800
tgctatctgc cgggggactg gctggaccag gcgggcgcgc ggatcgacgg gccggtgccg
4860
tcgccggagc tgtacacagt gatcctccgg ctgttgatg aggcggaacc ctattacgcg
4920
tcggcgcggg tgggtctggc ggatctgcc ccgcgctgc cctggtccat cgccgcgcgc
4980
ctacggatct atcgcgccat cgggctgcgc atccgaaga gcgggcccga ggcctatcgc
5040
cagcggatca gcacgtcaa ggctgccaag atcgccctgc tgggcgtcgc ggcctgggat
5100
gtcgcgcgat cacgcctgcc gggggcgggc gtgtcgcgc aggcctctg gaccggcgcc
5160
catcacgtct aggcgcgcgc ggcgtagggc agaaccggt ccagcagggc cgcgatttcc
5220
ggagcctgaa ggcgcttgct gcgcagcatc gcgtccagtt gggcgcggt ggcctcgtaa
5280

tgacgggaca cgttctgcag gtctgacacg gccagaaggc cgcgccgcgg gccgggggccc
5340

gcggcatcgc gaccggtatc cttgccaagc gccgcctggt cgcccacgac gtccagcagg
5400

tcgtcatagg actggaacac ggggccacgc tgacggccaa agtcgatcat ctgggtctgc
5460

tcctcggcgt cgaactcctt gatcacggcc agcatctcca gcccggcgat gaacagcacg
5520

ccggtcttca ggtcctgttc ctgttcgacc cccgcgccgt tcttgccgc gtgcaggctc
5580

aggctcctggc cggcgcacag gccctgcggc cccagggacc gcgacaggat ccgcaccagc
5640

tgcgcccga cctgcccga cgcgccgcgc gcaccggcca gcagggccat tgcctcggtg
5700

atcaggcgga tgccgccag cacggcacgg ctttcgccat gcgccacatg ggtcgcgggc
5760

cgccgcggc gcagccggc atcgtccatg cagggcaggc cgtcgaagat cagcgatgcg
5820

gcacgcacca tctcgaccgc gcaggcggcg tcgacgatcg tgctgcagac cccgcccgag
5880

gcctctgccg caagcagcat cagcatgccg cggaaccgcc tgcccagca cagcgcgcca
5940

tggctcatgg ccgcgccgag cggtcgcac acggcaccca atccctgggc gatctcctca
6000

agtctggtct gcagaagggt ggcgtggatc gggttgacgt ctgctctcat cagtgccttc
6060

gcgcttgggt tctgacctg cggaaggctc aggcggggc ggcaccccg gaccgctcat
6120

ccaccgtcaa cagtcccat gttggaacgg ttcacgccg attgcgagcc ttttcgacgg
6180

cgacgcgggg tcgcgcggca atttgtccaa caaggtcagt gga
6223

<210> 5
<211> 729
<212> DNA
<213> crtW gene

<400> 5
atgagcgcac atgccctgcc caaggcagat ctgaccgcca ccagcctgat cgtctcgggc
60

ggcatcatcg ccgcgtggct ggccctgcat gtgcatgcgc tgtggtttct ggacgcggcg
120

gcgcattcca tcctggcgat cgcgaatttc ctggggctga cctggctgtc ggtaggtctg
180

ttcttcatcg cgcattgacgc gatgcacggg tcggctgtgc cggggcgctcc gcgcggcaat
240

gcggcgatgg gccagctggt cctgtggctg tatgccggat ttctgtggcg caagatgac
300

gtcaagcaca tggcccatca ccgccatacc ggaaccgacg acgaccccgca ttctgaccat
360

ggcgggcccg tccgctggta cgcgcgcttc atcggcacct atttcggctg gcgcgagggg
420

ctgctgctgc ccgtcatcgt gacggctctat gcgctgatcc tgggggatcg ctggatgtac
480

gtggtcttct ggccgctgcc gtcgatcctg gcgtcgatcc agctgttcgt gttcggcacc
540

tggctgccgc accgccccgg ccacgacgcg ttcccggacc gccataatgc gcggctgtcg
600

cggatcagcg accccgtgtc gctgctgacc tgctttcact ttggtgggtta tcatcacgaa
660

caccacctgc acccgacggt gccttgggtg cgctgcca gcacccgcac caagggggac
720

accgcatga
729

<210> 6
<211> 242
<212> PRT
<213> crtW amino acid

<400> 6
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu 15
1 5 10
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
35 40 45
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Phe Ile Ala
50 55 60
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Gly Asn
65 70 75
80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp 95
 85 90
 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Thr Gly Thr
 100 105 110
 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
 115 120 125
 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
 130 135 140
 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
 145 150 155
 160
 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 175
 165 170
 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
 180 185 190
 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
 195 200 205
 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220
 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
 225 230 235
 240
 Thr Ala

<210> 7
 <211> 489
 <212> DNA
 <213> crtZ gene

<400> 7
 atgaccaatt tcctgacgt cgctgccacc gtgctggtga tggagttgac ggcctattcc
 60
 gtccaccggtt ggatcatgca cggccccctg ggctggggct ggcacaagtc ccaccacgag
 120
 gaacacgacc acgcgtgga aaagaacgac ctgtacggcc tggctctttgc ggtgatcgcc
 180
 acggtgctgt tcacggtggg ctggatctgg gcgccggtcc tgtggtggat cgctttgggc
 240
 atgaccgtct atgggctgat ctatttcgtc ctgcatgacg ggctgggttca tcagcgctgg
 300

ccgttccgct atatcccgcg caagggctat gcccgccgcc tgtatcaggc ccaccgcctg
360

caccacgcgg tcgagggacg cgaccattgc gtcagcttcg gcttcatcta tgcgccgccg
420

gtcgacaagc tgaagcagga cctgaagacg tcgggcgtgc tgcgggccga ggcgcaggag
480

cgcacgtga
489

<210> 8
<211> 162
<212> PRT
<213> crtZ amino acid

<400> 8
Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 15
1 5 10
Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
20 25 30
Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
35 40 45
Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe
50 55 60
Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly
65 70 75
80
Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val 95
85 90
His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
100 105 110
Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp
115 120 125
His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu
130 135 140
Lys Gln Asp Leu Lys Thr Ser Gly Val Leu Arg Ala Glu Ala Gln Glu
145 150 155
160
Arg Thr

<210> 9
<211> 1161
<212> DNA

<213> crtY gene

<400> 9

gtgaccatg acgtgctgct ggcaggggcg ggccttgca acgggctgat cgccctggcg
60

ctgcgcgcg cgcgggccga cctgcgggtg ctgctgctgg atcatgcggc gggaccgtca
120

gacggccata cctggtcctg ccacgacccc gatctgtcgc cgcactggct ggcgcggctg
180

aagcccctgc gccgcgcaa ctggcccgc caggaggtgc gctttcccg ccatgcccgg
240

cggttgcca ccggttacgg gtcgctggac gggcgggcgc tggcggatgc ggtggcccg
300

tcgggcgcg agatccgctg gaacagcgac atcgccctgc tggatgaaca gggggcgacg
360

ctgtcctgcg gcacccgat cgaggcgggc gcggtcctgg acgggcgcg cgcgacgcg
420

tcgcggcatc tgaccgtggg tttccagaaa ttctgggcg tcgagatcga gaccgactgc
480

ccccacggcg tgccccgcc gatgatcatg gacgcgaccg tcacccagca ggacgggtac
540

cgattcatct atctgctgcc cttctctccg acgcgcatcc tgatcgagga cactcgctat
600

tccgatggcg gcaatctgga cgacgacgcg ctggcggcgg cgtccacga ctatgcccg
660

cagcagggct ggaccggggc cgaggctccg cggaacgcg gcacccctgc cattgcgctg
720

gcccatacgc cggcgggctt ctgggccgat cacgcggagg ggcctgttcc cgtgggactg
780

cgcgcggggt tctttcaccg ggtcaccggc tattcgctgc cctatgcggc gcaggtggcg
840

gacgtggtgg cgggcctgtc cgggccgcc ggcaccgacg cgctgcgcg cgccatccgc
900

gattacgcga tcgaccgggc acgccgtgac cgctttctgc gcctgctgaa ccggatgctg
960

ttccgcggct gcgcgcccga ccggcgctat accctgctgc agcggttcta ccgcatgccg
1020

catggactga tcgaacggtt ctatgccggc cggctgagcg tggcggatca gctgcgcatc
1080

gtgaccggca agcctcccat tccccttggc acggccatcc gctgcctgcc cgaacgtccc

1140

ctgctgaagg aaaacgcatg a
1161

<210> 10
<211> 386
<212> PRT
<213> crtY amino acid

<400> 10
Val Thr His Asp Val Leu Leu Ala Gly Ala Gly Leu Ala Asn Gly Leu 15
1 5 10
Ile Ala Leu Ala Leu Arg Ala Ala Arg Pro Asp Leu Arg Val Leu Leu 30
20 25 30
Leu Asp His Ala Ala Gly Pro Ser Asp Gly His Thr Trp Ser Cys His 45
35 40 45
Asp Pro Asp Leu Ser Pro His Trp Leu Ala Arg Leu Lys Pro Leu Arg 60
50 55 60
Arg Ala Asn Trp Pro Asp Gln Glu Val Arg Phe Pro Arg His Ala Arg 75
65 70 75
80
Arg Leu Ala Thr Gly Tyr Gly Ser Leu Asp Gly Ala Ala Leu Ala Asp 95
85 90 95
Ala Val Ala Arg Ser Gly Ala Glu Ile Arg Trp Asn Ser Asp Ile Ala 110
100 105 110
Leu Leu Asp Glu Gln Gly Ala Thr Leu Ser Cys Gly Thr Arg Ile Glu 125
115 120 125
Ala Gly Ala Val Leu Asp Gly Arg Gly Ala Gln Pro Ser Arg His Leu 140
130 135 140
Thr Val Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Cys 155
145 150 155
160
Pro His Gly Val Pro Arg Pro Met Ile Met Asp Ala Thr Val Thr Gln 175
165 170 175
Gln Asp Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg 190
180 185 190
Ile Leu Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asn Leu Asp Asp 205
195 200 205
Asp Ala Leu Ala Ala Ala Ser His Asp Tyr Ala Arg Gln Gln Gly Trp 220
210 215 220
Thr Gly Ala Glu Val Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu 235
225 230 235

240

Ala His Asp Ala Ala Gly Phe Trp Ala Asp His Ala Glu Gly Pro Val 255
245 250

Pro Val Gly Leu Arg Ala Gly Phe Phe His Pro Val Thr Gly Tyr Ser 270
260 265

Leu Pro Tyr Ala Ala Gln Val Ala Asp Val Val Ala Gly Leu Ser Gly 285
275 280

Pro Pro Gly Thr Asp Ala Leu Arg Gly Ala Ile Arg Asp Tyr Ala Ile 300
290 295

Asp Arg Ala Arg Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu 315
305 310

Phe Arg Gly Cys Ala Pro Asp Arg Arg Tyr Thr Leu Leu Gln Arg Phe 335
325 330

Tyr Arg Met Pro His Gly Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu 350
340 345

Ser Val Ala Asp Gln Leu Arg Ile Val Thr Gly Lys Pro Pro Ile Pro 365
355 360

Leu Gly Thr Ala Ile Arg Cys Leu Pro Glu Arg Pro Leu Leu Lys Glu 380
370 375

Asn Ala 385

<210> 11
<211> 1506
<212> DNA
<213> crtI gene

<400> 11
atgaacgccc attcgccgc ggccaagacc gccatcgtga tcggcgagg ctttgccggg
60

ctggccctgg ccatccgcct gcagtcgcg gccatcgcca ccaccctggt cgaggcccg
120

gacaagccc gcgggcgcg ctatgtctgg cacgatcagg gccatgtctt cgacgcgggc
180

ccgaccgtca tcaccgacc cgatgcgctc aaggagctgt gggcgctgac cgggcaggac
240

atggcgcgcg acgtgacgt gatgccggtg tcgcccttct atcgactgat gtggccgggc
300

gggaaggtct tcgattacgt gaacgaggcc gatcagctgg agcgccagat cgcccagttc
360

aaccgagacg acctggaagg ataccgccgc ttccgtgatt acgaggagga ggtgtatcag
420
gagggctacg tcaagctggg caccgtgcc ttctcaagc tgggccagat gctcaaggcc
480
gcgcccgcgc tgatgaagct ggaggcctat aagtccgtcc atgccaaggt cgcgaccttc
540
atcaaggacc cctatctgcg gcaggcggtt tcgtatcaca cgctgctggt gggcgggaat
600
cccttctcga ccagctcgat ctatgcgctg atccacgcgc tggagcggcg cggcggggtc
660
tggttcgcca agggcggcac caaccagctg gtcgcgggca tggtcgcgct gttcgaacgg
720
cttggcgggc agatgatgct gaacgccaag gtcgccgga tcgagaccga gggcgcgcg
780
accacgggcg tcaccctggc ggacgggcg tctttaaggg ccgacatggt cgccagcaac
840
ggcgagctca tgcacaacta tcgcgacctg ctgggccaca cggcccgcg gcagagccgc
900
gcgaaatcgc tggaccgcaa gcgctggtcc atgtcgttgt tcgtgctgca tttcggctcg
960
cgcgaggcgc ccaaggacat cgcgcacac accatcctgt tcggcccccg ctacaggag
1020
ctggtcaacg agatcttcaa gggcccgaag ctggccgagg atttctcgct gtacctgcat
1080
tcgccctgca cgaccgatcc ggacatggcg cctccgggca tgtccacgca ttacgtgctg
1140
gccccgtgc cgcacatggg ccgcgccgag atcgattggg cggtcgaggg gccgcgctat
1200
gccgaccgca tcctggcgtc cctggaggag cggctgatcc cgaacctgc gcaccaacctg
1260
accacgacgc gcatcttcac gcccgccgat ttcgccagcg aactgaacgc ccatcacggc
1320
agcgcttct cggtcgagcc gatcctgacg caatccgcgt ggttccggcc gcacaaccgc
1380
gacaagacga tccgcaactt ctatctggtc ggcgcgggca cccatccggg cgcgggcatt
1440
ccgggctcg tgggctcggc caaggccacg gccaggtga tgctgtccga cctggcgggc
1500
gcatga

1506

<210> 12
<211> 501
<212> PRT
<213> crtI amino acid

<400> 12

Met Asn Ala His Ser Pro Ala Ala Lys Thr Ala Ile Val Ile Gly Ala 15
1 5 10

Gly Phe Gly Gly Leu Ala Leu Ala Ile Arg Leu Gln Ser Ala Gly Ile 30
20 25

Ala Thr Thr Leu Val Glu Ala Arg Asp Lys Pro Gly Gly Arg Ala Tyr 45
35 40

Val Trp His Asp Gln Gly His Val Phe Asp Ala Gly Pro Thr Val Ile 60
50 55

Thr Asp Pro Asp Ala Leu Lys Glu Leu Trp Ala Leu Thr Gly Gln Asp 75
65 70 80

Met Ala Arg Asp Val Thr Leu Met Pro Val Ser Pro Phe Tyr Arg Leu 95
85 90

Met Trp Pro Gly Gly Lys Val Phe Asp Tyr Val Asn Glu Ala Asp Gln 110
100 105

Leu Glu Arg Gln Ile Ala Gln Phe Asn Pro Asp Asp Leu Glu Gly Tyr 125
115 120

Arg Arg Phe Arg Asp Tyr Ala Glu Glu Val Tyr Gln Glu Gly Tyr Val 140
130 135

Lys Leu Gly Thr Val Pro Phe Leu Lys Leu Gly Gln Met Leu Lys Ala 155
145 150 160

Ala Pro Ala Leu Met Lys Leu Glu Ala Tyr Lys Ser Val His Ala Lys 175
165 170

Val Ala Thr Phe Ile Lys Asp Pro Tyr Leu Arg Gln Ala Phe Ser Tyr 190
180 185

His Thr Leu Leu Val Gly Gly Asn Pro Phe Ser Thr Ser Ser Ile Tyr 205
195 200

Ala Leu Ile His Ala Leu Glu Arg Arg Gly Gly Val Trp Phe Ala Lys 220
210 215

Gly Gly Thr Asn Gln Leu Val Ala Gly Met Val Ala Leu Phe Glu Arg 235
225 230 240

Leu Gly Gly Gln Met Met Leu Asn Ala Lys Val Ala Arg Ile Glu Thr

<400> 13
 atgagcgcgc tggctcctgac ctgcaccgag gcgatcacc aagggtcgca aagctttgcc
 60
 acggcggcca agctgatgcc gccgggcatc cgcgacgaca cggatgatgct ctatgcctgg
 120
 tgccgccacg cggatgacgt gatcgacggc caggccctgg gcagccgccc cgaggcggcg
 180
 aacgaccgcg aggcgcggct ggacggcctg cgcgtcgaca cgctggcggc cctgcagggc
 240
 gacggtccgg tgaccccgcc ctttgccgag ctgcgcgcgg tggcgcggcg gcatgatttc
 300
 ccgcaggcct ggcccatgga cctgatcgaa ggcttcgcga tggatgtcga ggcgcgcgac
 360
 tatcgacgcg tggatgacgt gctggaatat tcctatcacg tcgcaggcat cgtcggcgtg
 420
 atgatggccc gcgtgatggg cgtgcgcgac gatcctgtcc tggaccgcgc ctgcgacctg
 480
 gggctggcgt tccagctgac caacatcgcg cgcgacgtga tcgacgatgc gcgcacggg
 540
 cggatgctatc tgccggggga ctggctggac caggcggggc gcggatcga cgggccggcg
 600
 ccgtgcgcgg agctgtacac agtgatcctc cggctgttgg atgaggcgga accctattac
 660
 gcgtcggcgc ggggtgggtct ggccgatctg ccaccgcgct gcgcctggtc catcgccgcc
 720
 gcgctacgga tctatcgcg catcgggctg cgcacccgca agagcgggccc gcaggcctat
 780
 cgccagcgga tcagcacgtc caaggctgcc aagatcggcc tgctgggcgt cgggggctgg
 840
 gatgtcgcgc gatcacgcct gccggggggc ggctgtgcgc ggcagggcct ctggaccggg
 900
 ccgcatcacg tctag
 915

<210> 14
 <211> 304
 <212> PRT
 <213> crtB amino acid

<400> 14
 Met Ser Asp Leu Val Leu Thr Ser Thr Glu Ala Ile Thr Gln Gly Ser

1	5	10	15
Gln Ser Phe Ala Thr	Ala Ala Lys Leu Met	Pro Pro Gly Ile Arg Asp	
	20	25	30
Asp Thr Val Met Leu Tyr	Ala Trp Cys Arg His	Ala Asp Asp Val Ile	
	35	40	45
Asp Gly Gln Ala Leu Gly	Ser Arg Pro Glu Ala Val	Asn Asp Pro Gln	
	50	55	60
Ala Arg Leu Asp Gly Leu	Arg Val Asp Thr Leu	Ala Ala Leu Gln Gly	
	65	70	75
80			
Asp Gly Pro Val Thr	Pro Pro Phe Ala Ala	Leu Arg Ala Val Ala Arg	
	85	90	95
Arg His Asp Phe Pro Gln	Ala Trp Pro Met Asp	Leu Ile Glu Gly Phe	
	100	105	110
Ala Met Asp Val Glu Ala	Arg Asp Tyr Arg Thr	Leu Asp Asp Val Leu	
	115	120	125
Glu Tyr Ser Tyr His Val	Ala Gly Ile Val Gly	Val Met Met Ala Arg	
	130	135	140
Val Met Gly Val Arg Asp	Asp Pro Val Leu Asp	Arg Ala Cys Asp Leu	
	145	150	155
160			
Gly Leu Ala Phe Gln Leu	Thr Asn Ile Ala Arg	Asp Val Ile Asp Asp	
	165	170	175
Ala Arg Ile Gly Arg Cys	Tyr Leu Pro Gly Asp	Trp Leu Asp Gln Ala	
	180	185	190
Gly Ala Arg Ile Asp Gly	Pro Val Pro Ser Pro	Glu Leu Tyr Thr Val	
	195	200	205
Ile Leu Arg Leu Leu Asp	Glu Ala Glu Pro Tyr	Tyr Ala Ser Ala Arg	
	210	215	220
Val Gly Leu Ala Asp Leu	Pro Pro Arg Cys Ala	Trp Ser Ile Ala Ala	
	225	230	235
240			
Ala Leu Arg Ile Tyr Arg	Ala Ile Gly Leu Arg	Ile Arg Lys Ser Gly	
	245	250	255
Pro Gln Ala Tyr Arg Gln	Arg Ile Ser Thr Ser	Lys Ala Ala Lys Ile	
	260	265	270
Gly Leu Leu Gly Val Gly	Gly Trp Asp Val Ala	Arg Ser Arg Leu Pro	
	275	280	285
Gly Ala Gly Val Ser Arg	Gln Gly Leu Trp Thr	Arg Pro His His Val	
	290	295	300

<210> 15
<211> 882
<212> DNA
<213> crtE gene

<400> 15
atgagacgag acgtcaaccc gatccacgcc acccttctgc agaccagact tgaggagatc
60

gccacgggat tcggtgccgt gtcgcagccg ctccgcccgc ccatgagcca tggcgcgctg
120

tcgtcgggca ggcgggtccg cggcatgctg atgctgcttg cggcagaggc ctccggcggg
180

gtctgcgaca cgatcgtcga cgccgcctgc gcggtcgaga tggcgcgctg cgcacgctg
240

atcttcgacg acctgccctg catggacgat gccgggctgc gccgcggccg gcccgcgacc
300

catgtggcgc atggcgaaag ccgtgccgtg ctgggcggca tcgccctgat caccgaggca
360

atggccctgc tggccggtgc gcgcggcgcg tcgggcacgg tgcgggcgca gctggtgcg
420

atcctgtcgc ggtccctggg gccgcagggc ctgtgcgccg gccaggacct ggacctgcac
480

gcggccaaga acggcgcggg ggtcgaacag gaacaggacc tgaagaccg cgtgctgttc
540

atcgccgggc tggagatgct ggccgtgatc aaggagtctg acgccgagga gcagaccag
600

atgatcgaact ttggccgtca gctgggcccgc gtgttccagt cctatgacga cctgctggac
660

gtcgtgggcg accaggcggc gcttggcaag gataccggtc gcgatgccgc ggcccccg
720

ccgcggcgcg gccttctggc cgtgtcagac ctgcagaacg tgtcccgta ttacgaggcc
780

agccgcgccc aactggacgc gatgctgcgc agcaagcgcc ttcaggctcc ggaaatcgcg
840

gccctgctgg aacgggttct gccctacgcc gcgcgcgcct ag
882

<210> 16
<211> 293

```

<212>    PRT
<213>    crtE amino acid

<400>    16
Met Arg Arg Asp Val Asn Pro Ile His Ala Thr Leu Leu Gln Thr Arg    15
 1              5              10

Leu Glu Glu Ile Ala Gln Gly Phe Gly Ala Val Ser Gln Pro Leu Gly
      20              25              30

Ala Ala Met Ser His Gly Ala Leu Ser Ser Gly Arg Arg Phe Arg Gly
      35              40              45

Met Leu Met Leu Leu Ala Ala Glu Ala Ser Gly Gly Val Cys Asp Thr
 50              55              60

Ile Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu
 65              70              75
80

Ile Phe Asp Asp Leu Pro Cys Met Asp Asp Ala Gly Leu Arg Arg Gly
      85              90              95

Arg Pro Ala Thr His Val Ala His Gly Glu Ser Arg Ala Val Leu Gly
      100             105             110

Gly Ile Ala Leu Ile Thr Glu Ala Met Ala Leu Leu Ala Gly Ala Arg
      115             120             125

Gly Ala Ser Gly Thr Val Arg Ala Gln Leu Val Arg Ile Leu Ser Arg
      130             135             140

Ser Leu Gly Pro Gln Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His
145              150              155
160

Ala Ala Lys Asn Gly Ala Gly Val Glu Gln Glu Gln Asp Leu Lys Thr
      165             170             175

Gly Val Leu Phe Ile Ala Gly Leu Glu Met Leu Ala Val Ile Lys Glu
      180             185             190

Phe Asp Ala Glu Glu Gln Thr Gln Met Ile Asp Phe Gly Arg Gln Leu
      195             200             205

Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Val Gly Asp
      210             215             220

Gln Ala Ala Leu Gly Lys Asp Thr Gly Arg Asp Ala Ala Ala Pro Gly
225              230              235
240

Pro Arg Arg Gly Leu Leu Ala Val Ser Asp Leu Gln Asn Val Ser Arg
      245             250             255

His Tyr Glu Ala Ser Arg Ala Gln Leu Asp Ala Met Leu Arg Ser Lys
      260             265             270

```

Arg Leu Gln Ala Pro Glu Ile Ala Ala Leu Leu Glu Arg Val Leu Pro
275 280 285

Tyr Ala Ala Arg Ala
290

<210> 17
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer for crt gene

<400> 17
gttccacgac tggggcatc
19

<210> 18
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse primer for crt gene

<400> 18
tccactgacc ttgttgaca aattgccg
28